

SEQUENCE LISTING

<110> Meyers, Rachel A.
MacBeth, Kyle J.

<120> 14094, A NOVEL TRYPSIN FAMILY MEMBER AND
USES THEREFOR

<130> 10448-046002

<150> US 09/633,300

<151> 2000-08-08

<150> US 60/200,621

<151> 2000-04-28

<160> 13

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 2948

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (628)...(1986)

<221> misc_feature

<222> (1)...(2948)

<223> n = A,T,C or G

<400> 1

```

aagagttgca tatcgctcc catcaacaaa ctttcctgtg atttccanac aatgtatattt    60
gtttgtcaaa tccagttttc ttgtaaacad tggggggtaa ataacagagg tggcttatga    120
gtattttcttc cagggtaaaa agcaaaagaa ttccgggttt ctgtatcctt ttcacttact    180
gttaccactact ttgcctcgtc ttcaccctgt ccaaacaccg gtctccaatt tgcccttcag    240
agaacttaag tcaaggagag ttgaaattca caggccaggg cacatctttt atttatttca    300
ttatgttggc caacagaact tgattgtaaa taataataaa gaaatctgtt atatactttc    360
caaactccaa aaaaaaaccc gaattcagcc tggttaagtc caagctgaat tccgggtggg    420
ggaaggaccg ggcaccggac ggctcgggta ctttcgttct taattagggtc atgcccgtat    480
gagccaggaa agggctgtgt ttatgggaag ccagtaacac tgtggcctac tatctcttcc    540
gtggtgccat ctacattttt gggactcggg aattatgagg tagagggtgga ggcggagccg    600
gatgtcagag gtcttgaaat agtcacc atg ggg gaa aat gat ccg cct gct gtt    654
Met Gly Glu Asn Asp Pro Pro Ala Val

```

1

5

```

gaa gcc ccc ttc tca ttc cga tcg ctt ttt ggc ctt gat gat ttg aaa    702
Glu Ala Pro Phe Ser Phe Arg Ser Leu Phe Gly Leu Asp Asp Leu Lys
10                      15                      20                      25

```

```

ata agt cct gtt gca cca gat gca gat gct gtt gct gca cag atc ctg    750
Ile Ser Pro Val Ala Pro Asp Ala Asp Ala Val Ala Ala Gln Ile Leu
30                      35                      40

```

tca	ctg	ctg	cca	ttg	aag	ttt	ttt	cca	atc	atc	gtc	att	ggg	atc	att	798
Ser	Leu	Leu	Pro	Leu	Lys	Phe	Phe	Pro	Ile	Ile	Val	Ile	Gly	Ile	Ile	
			45				50						55			
gca	ttg	ata	tta	gca	ctg	gcc	att	ggg	ctg	ggc	atc	cac	ttc	gac	tgc	846
Ala	Leu	Ile	Leu	Ala	Leu	Ala	Ile	Gly	Leu	Gly	Ile	His	Phe	Asp	Cys	
			60				65						70			
tca	ggg	aag	tac	aga	tgt	cgc	tca	tcc	ttt	aag	tgt	atc	gag	ctg	ata	894
Ser	Gly	Lys	Tyr	Arg	Cys	Arg	Ser	Ser	Phe	Lys	Cys	Ile	Glu	Leu	Ile	
			75	80						85						
gct	cga	tgt	gac	gga	gtc	tcg	gat	tgc	aaa	gac	ggg	gag	gac	gag	tac	942
Ala	Arg	Cys	Asp	Gly	Val	Ser	Asp	Cys	Lys	Asp	Gly	Glu	Asp	Glu	Tyr	
90			95			100						105				
cgc	tgt	gtc	cgg	gtg	ggg	ggg	cag	aat	gcc	gtg	ctc	cag	gtg	ttc	aca	990
Arg	Cys	Val	Arg	Val	Gly	Gly	Gln	Asn	Ala	Val	Leu	Gln	Val	Phe	Thr	
			110			115						120				
gct	gct	tcg	tgg	aag	acc	atg	tgc	tcc	gat	gac	tgg	aag	ggg	cac	tac	1038
Ala	Ala	Ser	Trp	Lys	Thr	Met	Cys	Ser	Asp	Asp	Trp	Lys	Gly	His	Tyr	
			125			130						135				
gca	aat	gtt	gcc	tgt	gcc	caa	ctg	ggg	ttc	cca	agc	tat	gtg	agt	tca	1086
Ala	Asn	Val	Ala	Cys	Ala	Gln	Leu	Gly	Phe	Pro	Ser	Tyr	Val	Ser	Ser	
140						145						150				
gat	aac	ctc	aga	gtg	agc	tcg	ctg	gag	ggg	cag	ttc	cgg	gag	gag	ttt	1134
Asp	Asn	Leu	Arg	Val	Ser	Ser	Leu	Glu	Gly	Gln	Phe	Arg	Glu	Glu	Phe	
155						160						165				
gtg	tcc	atc	gat	cac	ctc	ttg	cca	gat	gac	aag	gtg	act	gca	tta	cac	1182
Val	Ser	Ile	Asp	His	Leu	Leu	Pro	Asp	Asp	Lys	Val	Thr	Ala	Leu	His	
170			175			180						185				
cac	tca	gta	tat	gtg	agg	gag	gga	tgt	gcc	tct	ggc	cac	gtg	gtt	acc	1230
His	Ser	Val	Tyr	Val	Arg	Glu	Gly	Cys	Ala	Ser	Gly	His	Val	Val	Thr	
			190			195						200				
ttg	cag	tgc	aca	gcc	tgt	ggg	cat	aga	agg	ggc	tac	agc	tca	cgc	atc	1278
Leu	Gln	Cys	Thr	Ala	Cys	Gly	His	Arg	Arg	Gly	Tyr	Ser	Ser	Arg	Ile	
			205			210						215				
gtg	ggg	gga	aac	atg	tcc	ttg	ctc	tcg	cag	tgg	ccc	tgg	cag	gcc	agc	1326
Val	Gly	Gly	Asn	Met	Ser	Leu	Leu	Ser	Gln	Trp	Pro	Trp	Gln	Ala	Ser	
220			225			230										
ctt	cag	ttc	cag	ggc	tac	cac	ctg	tgc	ggg	ggc	tct	gtc	atc	acg	ccc	1374
Leu	Gln	Phe	Gln	Gly	Tyr	His	Leu	Cys	Gly	Gly	Ser	Val	Ile	Thr	Pro	
235			240			245										
ctg	tgg	atc	atc	act	gct	gca	cac	tgt	gtt	tat	gac	ttg	tac	ctc	ccc	1422
Leu	Trp	Ile	Ile	Thr	Ala	Ala	His	Cys	Val	Tyr	Asp	Leu	Tyr	Leu	Pro	
250			255			260						265				
aag	tca	tgg	acc	atc	cag	gtg	ggg	cta	gtt	tcc	ctg	ttg	gac	aat	cca	1470

Lys	Ser	Trp	Thr	Ile	Gln	Val	Gly	Leu	Val	Ser	Leu	Leu	Asp	Asn	Pro				
				270								275							
gcc	cca	tcc	cac	ttg	gtg	gag	aag	att	gtc	tac	cac	agc	aag	tac	aag	1518			
Ala	Pro	Ser	His	Leu	Val	Glu	Lys	Ile	Val	Tyr	His	Ser	Lys	Tyr	Lys				
				285								290							
cca	aag	agg	ctg	ggc	aat	gac	atc	gcc	ctt	atg	aag	ctg	gcc	ggg	cca	1566			
Pro	Lys	Arg	Leu	Gly	Asn	Asp	Ile	Ala	Leu	Met	Lys	Leu	Ala	Gly	Pro				
				300								305							
ctc	acg	ttc	aat	gaa	atg	atc	cag	cct	gtg	tgc	ctg	ccc	aac	tct	gaa	1614			
Leu	Thr	Phe	Asn	Glu	Met	Ile	Gln	Pro	Val	Cys	Leu	Pro	Asn	Ser	Glu				
				315								320							
gag	aac	ttc	ccc	gat	gga	aaa	gtg	tgc	tgg	acg	tca	gga	tgg	ggg	gcc	1662			
Glu	Asn	Phe	Pro	Asp	Gly	Lys	Val	Cys	Trp	Thr	Ser	Gly	Trp	Gly	Ala				
				330								335							
aca	gag	gat	gga	ggc	gac	gcc	tcc	cct	gtc	ctg	aac	cac	gcg	gcc	gtc	1710			
Thr	Glu	Asp	Gly	Gly	Asp	Ala	Ser	Pro	Val	Leu	Asn	His	Ala	Ala	Val				
				350								355							
cct	ttg	att	tcc	aac	aag	atc	tgc	aac	cac	agg	gac	gtg	tac	ggc	ggc	1758			
Pro	Leu	Ile	Ser	Asn	Lys	Ile	Cys	Asn	His	Arg	Asp	Val	Tyr	Gly	Gly				
				365								370							
atc	atc	tcc	ccc	tcc	atg	ctc	tgc	gcg	ggc	tac	ctg	acg	ggc	ggc	gtg	1806			
Ile	Ile	Ser	Pro	Ser	Met	Leu	Cys	Ala	Gly	Tyr	Leu	Thr	Gly	Gly	Val				
				380								385							
gac	agc	tgc	cag	ggg	gac	agc	ggg	ggg	ccc	ctg	gtg	tgt	caa	gag	agg	1854			
Asp	Ser	Cys	Gln	Gly	Asp	Ser	Gly	Gly	Pro	Leu	Val	Cys	Gln	Glu	Arg				
				395								400							
agg	ctg	tgg	aag	tta	gtg	gga	gcg	acc	agc	ttt	ggc	atc	ggc	tgc	gca	1902			
Arg	Leu	Trp	Lys	Leu	Val	Gly	Ala	Thr	Ser	Phe	Gly	Ile	Gly	Cys	Ala				
				410								415							
gag	gtg	aac	aag	cct	ggg	gtg	tac	acc	cgt	gtc	acc	tcc	ttc	ctg	gac	1950			
Glu	Val	Asn	Lys	Pro	Gly	Val	Tyr	Thr	Arg	Val	Thr	Ser	Phe	Leu	Asp				
				430								435							
tg	atc	cac	gag	cag	atg	gag	aga	gac	cta	aaa	acc	tg	aaa	agg	gaa	1996			
Trp	Ile	His	Glu	Gln	Met	Glu	Arg	Asp	Leu	Lys	Thr								
				445								450							
ggggacaagt	agccac	ctga	gttc	ctgagg	tgatga	agac	agcccc	gatcc	tccc	ctggac	2056								
tcccgtag	gaacct	gcac	acgag	cagac	accctt	ggag	ctctgag	ttc	cggc	caccagt	2116								
agcaggcccc	aaagagg	ccac	ccttcc	atct	gattcc	cagca	caacctt	caa	gctg	cttttt	2176								
gttttttgtt	tttttg	agat	ggagt	ctctgc	tctgtt	gccc	agget	ggagt	gcagt	ggcga	2236								
aatccctgct	cactgc	agcc	tccgctt	cccc	tggtt	caagc	gattct	cttg	cctcag	cttc	2296								
cccagtagct	gggacc	acag	gtgccc	gcga	ccacaccca	ctaatt	ttttg	tatttt	tagt		2356								
agagacagg	tttcacc	atg	ttggcc	aggc	tgctct	caaa	cccc	gaacct	caaat	gatgt	2416								
gcctgcttca	gcctccc	aca	gtgctg	ggat	tacagg	cata	ggccacc	acg	cctagc	ctca	2476								
cgctcctttc	tgatctt	cac	taaga	aacaaa	agaag	cagca	acttg	caagg	gcggc	ctttc	2536								
ccactgggtcc	atctggt	tttt																	

```
<210> 2
<211> 453
<212> PRT
<213> Homo sapiens
```

Met 1	Gly	Glu	Asn	Asp 5	Pro	Pro	Ala	Val	Glu 10	Ala	Pro	Phe	Ser	Phe 15	Arg
Ser	Leu	Phe	Gly 20	Leu	Asp	Asp	Leu	Lys 25	Ile	Ser	Pro	Val	Ala 30	Pro	Asp
Ala	Asp	Ala 35	Val	Ala	Ala	Gln	Ile 40	Leu	Ser	Leu	Leu	Pro 45	Leu	Lys	Phe
Phe	Pro 50	Ile	Ile	Val	Ile	Gly 55	Ile	Ile	Ala	Leu	Ile 60	Leu	Ala	Leu	Ala
Ile 65	Gly	Leu	Gly	Ile	His 70	Phe	Asp	Cys	Ser	Gly 75	Lys	Tyr	Arg	Cys	Arg
Ser	Ser	Phe	Lys	Cys 85	Ile	Glu	Leu	Ile	Ala 90	Arg	Cys	Asp	Gly	Val	Ser
Asp	Cys	Lys	Asp 100	Gly	Glu	Asp	Glu	Tyr 105	Arg	Cys	Val	Arg	Val 110	Gly	Gly
Gln	Asn 115	Ala	Val	Leu	Gln	Val	Phe 120	Thr	Ala	Ala	Ser	Trp 125	Lys	Thr	Met
Cys	Ser 130	Asp	Asp	Trp	Lys	Gly 135	His	Tyr	Ala	Asn 140	Val	Ala	Cys	Ala	Gln
Leu 145	Gly	Phe	Pro	Ser	Tyr 150	Val	Ser	Ser	Asp	Asn 155	Leu	Arg	Val	Ser	Ser
Leu	Glu	Gly	Gln	Phe 165	Arg	Glu	Glu	Phe	Val 170	Ser	Ile	Asp	His 175	Leu	Leu
Pro	Asp	Asp	Lys 180	Val	Thr	Ala	Leu	His 185	His	Ser	Val	Tyr 190	Val	Arg	Glu
Gly	Cys 195	Ala	Ser	Gly	His	Val	Val 200	Thr	Leu	Gln	Cys	Thr 205	Ala	Cys	Gly
His 210	Arg	Arg	Gly	Tyr	Ser	Ser 215	Arg	Ile	Val	Gly 220	Gly	Asn	Met	Ser	Leu
Leu 225	Ser	Gln	Trp	Pro	Trp 230	Gln	Ala	Ser	Leu	Gln 235	Phe	Gln	Gly	Tyr	His
Leu	Cys	Gly	Gly	Ser 245	Val	Ile	Thr	Pro	Leu 250	Trp	Ile	Ile	Thr 255	Ala	Ala
His	Cys	Val	Tyr 260	Asp	Leu	Tyr	Leu	Pro 265	Lys	Ser	Trp	Thr 270	Ile	Gln	Val
Gly	Leu 275	Val	Ser	Leu	Leu	Asp	Asn 280	Pro	Ala	Pro	Ser	His 285	Leu	Val	Glu
Lys 290	Ile	Val	Tyr	His	Ser	Lys 295	Tyr	Lys	Pro	Lys	Arg 300	Leu	Gly	Asn	Asp
Ile 305	Ala	Leu	Met	Lys	Leu 310	Ala	Gly	Pro	Leu	Thr 315	Phe	Asn	Glu	Met	Ile
Gln	Pro	Val	Cys 325	Leu	Pro	Asn	Ser	Glu	Glu 330	Asn	Phe	Pro	Asp 335	Gly	Lys
Val	Cys	Trp	Thr 340	Ser	Gly	Trp	Gly	Ala 345	Thr	Glu	Asp	Gly	Gly 350	Asp	Ala

```
<210> 3
<211> 1362
<212> DNA
<213> Homo sapiens
```

```
<210> 4
<211> 260
<212> PRT
<213> Artificial Sequence
```

<400> 4
Ile Val Gly Gly Arg Glu Ala Gln Pro Gly Ser Phe Gly Ser Pro Trp
1 5 10 15
Gln Val Ser Leu Gln Val Arg Ser Gly Gly Gly Ser Arg Lys His Phe

20 25 30
 Cys Gly Gly Ser Leu Ile Ser Glu Asn Trp Val Leu Thr Ala Ala His
 35 40 45
 Cys Val Ser Gly Ala Ala Ser Ala Pro Ala Ser Ser Val Arg Val Ser
 50 55 60
 Leu Ser Arg Val Arg Leu Gly Glu His Asn Leu Ser Leu Thr Glu Gly
 65 70 75 80
 Thr Glu Gln Lys Phe Asp Val Lys Lys Thr Ile Ile Val His Pro Asn
 85 90 95
 Tyr Asn Pro Asp Thr Leu Asp Asn Gly Ala Tyr Asp Asn Asp Ile Ala
 100 105 110
 Leu Leu Lys Leu Lys Ser Pro Gly Val Thr Leu Gly Asp Thr Val Arg
 115 120 125
 Pro Ile Cys Leu Pro Ser Ala Ser Ser Asp Leu Pro Val Gly Thr Thr
 130 135 140
 Cys Thr Val Ser Gly Trp Gly Arg Arg Pro Thr Lys Asn Leu Gly Leu
 145 150 155 160
 Ser Asp Thr Leu Gln Glu Val Val Val Pro Val Val Ser Arg Glu Thr
 165 170 175
 Cys Arg Ser Ala Tyr Glu Tyr Gly Gly Thr Asp Asp Lys Val Glu Phe
 180 185 190
 Val Thr Asp Asn Met Ile Cys Ala Gly Ala Leu Gly Gly Lys Asp Ala
 195 200 205
 Cys Gln Gly Asp Ser Gly Gly Pro Leu Val Cys Ser Asp Gly Asn Arg
 210 215 220
 Asp Gly Arg Trp Glu Leu Val Gly Ile Val Ser Trp Gly Ser Tyr Gly
 225 230 235 240
 Cys Ala Arg Gly Asn Lys Pro Gly Val Tyr Thr Arg Val Ser Ser Tyr
 245 250 255
 Leu Asp Trp Ile
 260

<210> 5
 <211> 226
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> consensus sequence

<400> 5
 Arg Ile Val Gly Gly Ser Glu Ala Lys Ile Gly Ser Phe Pro Trp Gln
 1 5 10 15
 Val Ser Leu Gln Cys Gly Gly Ser Leu Ile Ser Pro Arg Trp Val Leu
 20 25 30
 Thr Ala Ala His Cys Arg Val Arg Leu Gly Ser His Asp Leu Ser Ser
 35 40 45
 Gly Glu Glu Thr Glu Gly Gly Pro Arg Leu Asp Ser Pro Gly Gly Gln
 50 55 60
 Val Ile Lys Val Ser Lys Ile Ile Glu Val His Pro Asn Tyr Asn Asn
 65 70 75 80
 Asp Ile Ala Leu Leu Lys Leu Lys Glu Pro Val Thr Leu Ser Asp Ser
 85 90 95
 Asn Thr Val Arg Pro Ile Cys Leu Pro Ser Ser Asn Glu Ile Lys Thr
 100 105 110
 Ser Glu Gly Asn Thr Val Pro Ala Gly Thr Thr Cys Thr Val Ser Gly
 115 120 125

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100

Trp Gly Arg Thr Ser Glu Gly Pro Glu Glu Ser Gly Gly Gly Ser Leu
 130 135 140
 Pro Asp Val Leu Gln Glu Val Asn Val Pro Ile Val Ser Asn Glu Thr
 145 150 155 160
 Cys Arg Met Leu Cys Ala Gly Tyr Leu Glu Gly Gly Asn Thr Pro Gly
 165 170 175
 Gly Lys Asp Ala Cys Gln Gly Asp Ser Gly Gly Pro Leu Val Cys Val
 180 185 190
 Leu Val Gly Ile Val Ser Trp Gly Ser Ser Ser Leu Tyr Gly Cys Ala
 195 200 205
 Arg Pro Asn Lys Pro Gly Val Tyr Thr Arg Val Ser Ser Tyr Leu Asp
 210 215 220
 Trp Ile
 225

<210> 6
 <211> 43
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> consensus sequence

<400> 6
 Ser Thr Cys Gly Gly Pro Asp Glu Phe Gln Cys Gly Ser Gly Arg Arg
 1 5 10 15
 Cys Ile Pro Arg Ser Trp Val Cys Asp Gly Asp Pro Asp Cys Glu Asp
 20 25 30
 Gly Ser Asp Glu Ser Leu Glu Asn Cys Ala Ala
 35 40

<210> 7
 <211> 113
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> consensus sequence

<400> 7
 Val Gly Gly Ser Ser Arg Cys Glu Gly Arg Val Glu Val Arg His Asp
 1 5 10 15
 Gly Ser Lys Trp Gly Thr Val Cys Asp Ser Ser Trp Ser Leu Arg Asp
 20 25 30
 Ala Asn Val Asp Pro Gln Ala Ser Lys Val Cys Arg Gln Leu Gly Cys
 35 40 45
 Gly Gly Ala Val Ser Leu Leu Gly Pro Tyr Phe Ser Glu Gly Gly Gly
 50 55 60
 Pro Ala Gly Gln Arg Glu Ile Trp Leu Asp Gly Val Asn Cys Ser Gly
 65 70 75 80
 Asn Glu Thr Ser Leu Ser Gln Cys Pro Val Arg Val Thr Pro Pro Gly
 85 90 95
 Leu Ser Arg Gln Cys Ser His Asp Gly Glu Asp Ala Gly Val Val Cys
 100 105 110
 Ser

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

<210> 8
 <211> 5
 <212> PRT
 <213> Homo sapiens

<400> 8
 Arg Ile Val Gly Gly
 1 5

<210> 9
 <211> 5
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> motif

<221> VARIANT
 <222> 2
 <223> Xaa = Asp or Glu

<221> VARIANT
 <222> 5
 <223> Xaa = Gly or Ser

<400> 9
 Gly Xaa Ser Gly Xaa
 1 5

<210> 10
 <211> 6
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> motif

<221> VARIANT
 <222> 1
 <223> Xaa = Leu, Ile, Val, or Met

<221> VARIANT
 <222> 2
 <223> Xaa = Ser or Thr

<221> VARIANT
 <222> 4
 <223> Xaa = Ser, Thr, Ala, or Gly

<400> 10
 Xaa Xaa Ala Xaa His Cys
 1 5

<210> 11
 <211> 2951
 <212> DNA
 <213> Homo sapiens

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

<220>
 <221> CDS
 <222> (628)...(1989)

<221> misc_feature
 <222> (1)...(2951)
 <223> n = A,T,C or G

<400> 11
 aagagttgca tatcgctccc catcaacaaa ctttcctgtg atttccanac aatgtatttt 60
 gtttgtcaaa tccagttttc ttgtaaaagat tgggggggtaa ataacagagg tggccttatga 120
 gtattttcttc cagggtaaaa agcaaaaagaa ttccggtttt ctgtatcctt ttcacttact 180
 gttaccctact ttgcctcgtc ttcaccctgt ccaaacaccg gtctccaatt tgcccttcag 240
 agaacttaag tcaaggagag ttgaaattca caggccaggg cacatctttt atttattttca 300
 ttatgttggc caacagaact tgattgtaaa taataataaa gaaatctgtt atatactttc 360
 caaactccaa aaaaaaaccg gaattcagcc tgggttaagtc caagctgaat tccgggtggg 420
 ggaaggaccg ggcaccggac ggctcgggta ctttcgttct taattagggtc atgcccgtat 480
 gagccaggaa agggctgtgt ttatgggaag ccagtaacac tgtggcctac tatctcttcc 540
 gtggtgccat ctacattttt gggactcggg aattatgagg tagagggtgga ggcggagccg 600
 gatgtcagag gtcctgaaat agtcacc atg ggg gaa aat gat ccg cct gct gtt 654
 Met Gly Glu Asn Asp Pro Pro Ala Val
 1 5

gaa gcc ccc ttc tca ttc cga tcg ctt ttt ggc ctt gat gat ttg aaa 702
 Glu Ala Pro Phe Ser Phe Arg Ser Leu Phe Gly Leu Asp Asp Leu Lys
 10 15 20 25

ata agt cct gtt gca cca gat gca gat gct gtt gct gca cag atc ctg 750
 Ile Ser Pro Val Ala Pro Asp Ala Asp Ala Val Ala Ala Gln Ile Leu
 30 35 40

tca ctg ctg cca ttg aag ttt ttt cca atc atc gtc att ggg atc att 798
 Ser Leu Leu Pro Leu Lys Phe Phe Pro Ile Ile Val Ile Gly Ile Ile
 45 50 55

gca ttg ata tta gca ctg gcc att ggt ctg ggc atc cac ttc gac tgc 846
 Ala Leu Ile Leu Ala Leu Ala Ile Gly Leu Gly Ile His Phe Asp Cys
 60 65 70

tca ggg aag tac aga tgt cgc tca tcc ttt aag tgt atc gag ctg ata 894
 Ser Gly Lys Tyr Arg Cys Arg Ser Ser Phe Lys Cys Ile Glu Leu Ile
 75 80 85

gct cga tgt gac gga gtc tcg gat tgc aaa gac ggg gag gac gag tac 942
 Ala Arg Cys Asp Gly Val Ser Asp Cys Lys Asp Gly Glu Asp Glu Tyr
 90 95 100 105

cgc tgt gtc cgg gtg ggt ggt cag aat gcc gtg ctc cag gtg ttc aca 990
 Arg Cys Val Arg Val Gly Gly Gln Asn Ala Val Leu Gln Val Phe Thr
 110 115 120

gct gct tcg tgg aag acc atg tgc tcc gat gac tgg aag ggt cac tac 1038
 Ala Ala Ser Trp Lys Thr Met Cys Ser Asp Asp Trp Lys Gly His Tyr
 125 130 135

gca aat gtt gcc tgt gcc caa ctg ggt ttc cca agc tat gtg agt tca 1086

Ala	Asn	Val	Ala	Cys	Ala	Gln	Leu	Gly	Phe	Pro	Ser	Tyr	Val	Ser	Ser		
140						145						150					
gat	aac	ctc	aga	gtg	agc	tcg	ctg	gag	ggg	cag	ttc	cgg	gag	gag	ttt	1134	
Asp	Asn	Leu	Arg	Val	Ser	Ser	Leu	Glu	Gly	Gln	Phe	Arg	Glu	Glu	Phe		
155						160					165						
gtg	tcc	atc	gat	cac	ctc	ttg	cca	gat	gac	aag	gtg	act	gca	tta	cac	1182	
Val	Ser	Ile	Asp	His	Leu	Leu	Pro	Asp	Asp	Lys	Val	Thr	Ala	Leu	His		
170					175					180					185		
cac	tca	gta	tat	gtg	agg	gag	gga	tgt	gcc	tct	ggc	cac	gtg	gtt	acc	1230	
His	Ser	Val	Tyr	Val	Arg	Glu	Gly	Cys	Ala	Ser	Gly	His	Val	Val	Thr		
				190					195					200			
ttg	cag	tgc	aca	gcc	tgt	ggg	cat	aga	agg	ggc	tac	agc	tca	cgc	atc	1278	
Leu	Gln	Cys	Thr	Ala	Cys	Gly	His	Arg	Arg	Gly	Tyr	Ser	Ser	Arg	Ile		
			205					210					215				
gtg	ggg	gga	aac	atg	tcc	ttg	ctc	tcg	cag	tgg	ccc	tgg	cag	gcc	agc	1326	
Val	Gly	Gly	Asn	Met	Ser	Leu	Leu	Ser	Gln	Trp	Pro	Trp	Gln	Ala	Ser		
			220				225					230					
ctt	cag	ttc	cag	ggc	tac	cac	ctg	tgc	ggg	ggc	tct	gtc	atc	acg	ccc	1374	
Leu	Gln	Phe	Gln	Gly	Tyr	His	Leu	Cys	Gly	Gly	Ser	Val	Ile	Thr	Pro		
235						240					245						
ctg	tgg	atc	atc	act	gct	gca	cac	tgt	gtt	tat	gac	ttg	tac	ctc	ccc	1422	
Leu	Trp	Ile	Ile	Thr	Ala	Ala	His	Cys	Val	Tyr	Asp	Leu	Tyr	Leu	Pro		
250					255					260					265		
aag	tca	tgg	acc	atc	cag	gtg	ggg	cta	gtt	tcc	ctg	ttg	gac	aat	cca	1470	
Lys	Ser	Trp	Thr	Ile	Gln	Val	Gly	Leu	Val	Ser	Leu	Leu	Asp	Asn	Pro		
				270					275					280			
gcc	cca	tcc	cac	ttg	gtg	gag	aag	att	gtc	tac	cac	agc	aag	tac	aag	1518	
Ala	Pro	Ser	His	Leu	Val	Glu	Lys	Ile	Val	Tyr	His	Ser	Lys	Tyr	Lys		
			285					290					295				
cca	aag	agg	ctg	ggc	aat	gac	atc	gcc	ctt	atg	aag	ctg	gcc	ggg	cca	1566	
Pro	Lys	Arg	Leu	Gly	Asn	Asp	Ile	Ala	Leu	Met	Lys	Leu	Ala	Gly	Pro		
300						305					310						
ctc	acg	ttc	aat	gaa	atg	atc	cag	cct	gtg	tgc	ctg	ccc	aac	tct	gaa	1614	
Leu	Thr	Phe	Asn	Glu	Met	Ile	Gln	Pro	Val	Cys	Leu	Pro	Asn	Ser	Glu		
315						320					325						
gag	aac	ttc	ccc	gat	gga	aaa	gtg	tgc	tgg	acg	tca	gga	tgg	ggg	gcc	1662	
Glu	Asn	Phe	Pro	Asp	Gly	Lys	Val	Cys	Trp	Thr	Ser	Gly	Trp	Gly	Ala		
330					335					340					345		
aca	gag	gat	gga	gca	ggg	gac	gcc	tcc	cct	gtc	ctg	aac	cac	gcg	gcc	1710	
Thr	Glu	Asp	Gly	Ala	Gly	Asp	Ala	Ser	Pro	Val	Leu	Asn	His	Ala	Ala		
			350					355					360				
gtc	cct	ttg	att	tcc	aac	aag	atc	tgc	aac	cac	agg	gac	gtg	tac	ggg		

365	370	375	
ggc atc atc tcc ccc tcc atg ctc tgc gcg ggc tac ctg acg ggt ggc			1806
Gly Ile Ile Ser Pro Ser Met Leu Cys Ala Gly Tyr Leu Thr Gly Gly			
380	385	390	
gtg gac agc tgc cag ggg gac agc ggg ggg ccc ctg gtg tgt caa gag			1854
Val Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro Leu Val Cys Gln Glu			
395	400	405	
agg agg ctg tgg aag tta gtg gga gcg acc agc ttt ggc atc ggc tgc			1902
Arg Arg Leu Trp Lys Leu Val Gly Ala Thr Ser Phe Gly Ile Gly Cys			
410	415	420	425
gca gag gtg aac aag cct ggg gtg tac acc cgt gtc acc tcc ttc ctg			1950
Ala Glu Val Asn Lys Pro Gly Val Tyr Thr Arg Val Thr Ser Phe Leu			
430	435	440	
gac tgg atc cac gag cag atg gag aga gac cta aaa acc tgaagaggaa			1999
Asp Trp Ile His Glu Gln Met Glu Arg Asp Leu Lys Thr			
445	450		
ggggacaagt agccacctga gttcctgagg tgatgaagac agcccgatcc tcccctggac			2059
tcccgtgtag gaacctgcac acgagcagac acccttggag ctctgagttc cggcaccagt			2119
agcaggcccg aaagaggcac ccttccatct gattccagca caaccttcaa gctgcttttt			2179
gttttttggt tttttgagat ggagtctcgc tctgttgccc aggctggagt gcagtggcga			2239
aatccctgct cactgcagcc tccgcttccc tggttcaagc gattctcttg cctcagcttc			2299
cccagtagct gggaccacag gtgcccgcc ccacacccaa ctaatttttg tatttttagt			2359
agagacaggg tttcaccatg ttggccaggc tgctctcaaa cccctgacct caaatgatgt			2419
gcctgcttca gcctcccaca gtgctgggat tacaggcatg ggccaccacg cctagcctca			2479
cgctcctttc tgatcttcac taagaacaaa agaagcagca acttgcaagg gcggcctttc			2539
ccactggtcc atctgggttt ctctccaggg gtcttgcaaa attcctgacg agataagcag			2599
ttatgtgacc tcacgtgcaa agccaccaac agccactcag aaaagacgca ccagcccaga			2659
agtgcagaac tgcagtcact gcacgttttc atctctaggg accagaacca aaccaccct			2719
ttctacttcc aagacttatt ttcacatgtg gggagggttaa tctaggaatg actcggttaa			2779
ggcctatttt catgatttct ttgtagcatt tgggtgcttga cgtattattg tcctttgatt			2839
ccaaataata tgtttccttc cctcatwraa maaaaaaaaa aaaaaaaarr rmrrssgcta			2899
vavmarktta gagaaaaaac ctaccacrc cttccccctg aamctraaam ya			2951

<210> 12

<211> 454

<212> PRT

<213> Homo sapiens

<400> 12

Met Gly Glu Asn Asp Pro Pro Ala Val Glu Ala Pro Phe Ser Phe Arg			
1	5	10	15
Ser Leu Phe Gly Leu Asp Asp Leu Lys Ile Ser Pro Val Ala Pro Asp			
20	25	30	
Ala Asp Ala Val Ala Ala Gln Ile Leu Ser Leu Leu Pro Leu Lys Phe			
35	40	45	
Phe Pro Ile Ile Val Ile Gly Ile Ile Ala Leu Ile Leu Ala Leu Ala			
50	55	60	
Ile Gly Leu Gly Ile His Phe Asp Cys Ser Gly Lys Tyr Arg Cys Arg			
65	70	75	80
Ser Ser Phe Lys Cys Ile Glu Leu Ile Ala Arg Cys Asp Gly Val Ser			
85	90	95	

Asp Cys Lys Asp Gly Glu Asp Glu Tyr Arg Cys Val Arg Val Gly Gly
 100 105 110
 Gln Asn Ala Val Leu Gln Val Phe Thr Ala Ala Ser Trp Lys Thr Met
 115 120 125
 Cys Ser Asp Asp Trp Lys Gly His Tyr Ala Asn Val Ala Cys Ala Gln
 130 135 140
 Leu Gly Phe Pro Ser Tyr Val Ser Ser Asp Asn Leu Arg Val Ser Ser
 145 150 155 160
 Leu Glu Gly Gln Phe Arg Glu Glu Phe Val Ser Ile Asp His Leu Leu
 165 170 175
 Pro Asp Asp Lys Val Thr Ala Leu His His Ser Val Tyr Val Arg Glu
 180 185 190
 Gly Cys Ala Ser Gly His Val Val Thr Leu Gln Cys Thr Ala Cys Gly
 195 200 205
 His Arg Arg Gly Tyr Ser Ser Arg Ile Val Gly Gly Asn Met Ser Leu
 210 215 220
 Leu Ser Gln Trp Pro Trp Gln Ala Ser Leu Gln Phe Gln Gly Tyr His
 225 230 235 240
 Leu Cys Gly Gly Ser Val Ile Thr Pro Leu Trp Ile Ile Thr Ala Ala
 245 250 255
 His Cys Val Tyr Asp Leu Tyr Leu Pro Lys Ser Trp Thr Ile Gln Val
 260 265 270
 Gly Leu Val Ser Leu Leu Asp Asn Pro Ala Pro Ser His Leu Val Glu
 275 280 285
 Lys Ile Val Tyr His Ser Lys Tyr Lys Pro Lys Arg Leu Gly Asn Asp
 290 295 300
 Ile Ala Leu Met Lys Leu Ala Gly Pro Leu Thr Phe Asn Glu Met Ile
 305 310 315 320
 Gln Pro Val Cys Leu Pro Asn Ser Glu Glu Asn Phe Pro Asp Gly Lys
 325 330 335
 Val Cys Trp Thr Ser Gly Trp Gly Ala Thr Glu Asp Gly Ala Gly Asp
 340 345 350
 Ala Ser Pro Val Leu Asn His Ala Ala Val Pro Leu Ile Ser Asn Lys
 355 360 365
 Ile Cys Asn His Arg Asp Val Tyr Gly Gly Ile Ile Ser Pro Ser Met
 370 375 380
 Leu Cys Ala Gly Tyr Leu Thr Gly Gly Val Asp Ser Cys Gln Gly Asp
 385 390 395 400
 Ser Gly Gly Pro Leu Val Cys Gln Glu Arg Arg Leu Trp Lys Leu Val
 405 410 415
 Gly Ala Thr Ser Phe Gly Ile Gly Cys Ala Glu Val Asn Lys Pro Gly
 420 425 430
 Val Tyr Thr Arg Val Thr Ser Phe Leu Asp Trp Ile His Glu Gln Met
 435 440 445
 Glu Arg Asp Leu Lys Thr
 450

<210> 13

<211> 1365

<212> DNA

<213> Homo sapiens

<400> 13

atgggggaaa atgatccgcc tgctgttgaa gcccccttct cattccgatac gctttttggc	60
cttgatgatt tgaaaataag tctgttgca ccagatgcag atgctgttgc tgcacagatac	120
ctgtcactgc tgccattgaa gttttttcca atcatcgtca ttgggatcat tgcattgata	180
ttagcactgg ccattggtct gggcatccac ttcgactgct cagggaagta cagatgtcgc	240

tcaccccttta	agtgtatcga	gctgatagct	cgatgtgacg	gagtctcggg	ttgcaaagac	300
ggggaggacg	agtaccgctg	tgtccgggtg	ggtgggcaga	atgccgtgct	ccagggtgttc	360
acagctgctt	cgtggaagac	catgtgctcc	gatgactgga	agggtcacta	cgcaaatggt	420
gcctgtgccc	aactgggttt	cccaagctat	gtgagttcag	ataacctcag	agtgagctcg	480
ctggaggggc	agttccggga	ggagtttgtg	tccatcgatc	acctcttgcc	agatgacaag	540
gtgactgcat	tacaccactc	agtatatgtg	agggagggat	gtgcctctgg	ccacgtgggt	600
accttgccgt	gcacagcctg	tggatcataga	aggggctaca	gctcacgcat	cgtgggtgga	660
aacatgtcct	tgctctcgca	gtggccctgg	caggccagcc	ttcagttcca	gggctaccac	720
ctgtgcgggg	gctctgtcat	cacgcccctg	tggatcatca	ctgctgcaca	ctgtgtttat	780
gacttgtagc	tccccaagtc	atggaccatc	cagggtgggtc	tagtttccct	gttggacaat	840
ccagcccat	cccacttggt	ggagaagatt	gtctaccaca	gcaagtacaa	gccaaagagg	900
ctgggcaatg	acatcgccct	tatgaagctg	gccggggccac	tcacgttcaa	tgaaatgac	960
cagcctgtgt	gcctgcccac	ctctgaagag	aacttccccg	atggaaaagt	gtgctggacg	1020
tcaggatggg	gggcccacaga	ggatggagca	ggtgacgcct	cccctgtcct	gaaccacgcg	1080
gccgtccctt	tgatttccaa	caagatctgc	aaccacaggg	acgtgtacgg	tggcatcatc	1140
tccccctcca	tgctctgcgc	gggctacctg	acgggtggcg	tggacagctg	ccagggggac	1200
agcggggggc	ccctgggtgtg	tcaagagagg	aggctgtgga	agttagtggg	agcgaccagc	1260
tttggcatcg	gctgcgcaga	ggtgaacaag	cctgggggtgt	acaccctgtg	cacctccttc	1320
ctggactgga	tccacgagca	gatggagaga	gacctaaaaa	cctga		1365

09465.09465